音节的本质和元辅音性质新说*

江荻

江苏师范大学语言科学与艺术学院，中国社会科学院民族学与人类学研究所

摘要：人们长期把音节作为一级语音单位，却不知是个假象；元音和辅音一直是语音学最坚实的单位，却不知实际是字母的冒名假借；而因之形成元辅音构成音节的历史错案也无人洞察。文章分析了音节没有语言学地位的原因，叙述了元音和辅音身世不明的困惑，并从历时角度剖析音节与字母的形成渊源，以及从实验探索言语产生的单元形式，从共时角度揭示出发音时间结构对音节本质的规定。在此基础上，本文以颠覆式的思路再定义音节，重塑元辅音的真实地位和价值。文章最后提出，词的音节序列是在久远的历史演化中逐步形成的，呈现为复杂的词音节结构和有限的纯音节形式，据此可提出一套最简音节分析方案和依据历史演变过程的切分原则。

关键词：音节定义，元音音节，辅音音节，字母，切分音节

当代语音学中，元音音段和辅音音段是语音学和整个语言学最基本的建构单位，同时又是音节的直接构成单元，即元音和辅音以各种可能的组合方式构成音节（暂不论及韵律要素）。可是，不为人所知的是，语音学史上似乎从来没有人论证语言中为什么有元音和辅音，我们现有的从语流中切分元音和辅音操作是先有具体元辅音实体概念，再论述其性质、特征以及分类的理由。此外，音节也是语音系统中令人迷惑的单位，既无法定性也说不清其价值。针对以上问题，笔者梳理了元辅音产生的历史渊源、核查了音节的真实价值，发现音节、元音和辅音的本质和特征已在多种处理方法中获得认知突破和局部性解决，并呈现出他们的本质和相互关系。这些进展是近百年来语音学和音系学领域的重要成就，包括来自传统描写领域的深化、语音结构分析、语音心理认知（言语产生）、非线性音系学理论、以及语音声学实验等。可惜这些新知一直被强大的语音学传统掩蔽，学界习焉不察。本文欲深挖各领域取得的相关研究成果，据此重新定位元辅音性质和界定音节，企图为建构新的语音学体系铺垫基础。

1 音节的地位

1.1 音节从未获得语言学地位

语言研究史上，音节一直处于一个奇怪的地位：各种语言学论著或者教材没有不提及音节概念的，可又从来未把音节作为一级不可或缺的实体单位对待。更可叹的是，很少有研究者清醒意识到这一点，只是自认为音节理所当然是一级重要的语音单位。这件事情无论是 20 世纪以前的传统语言学，还是 20 世纪以来的索绪尔结构主义或者以音位学著称的布拉格学派，还有布隆菲尔德领衔的美国描写结构主义，使用“音节”术语是那么顺手那么随意，却没有给音节价值一个子丑寅卯的说道。这其中一定蕴含了某种特别的原因和导向。不妨观察一下美国结构语言学大师布隆菲尔德详尽描述的一个案例（布隆菲尔德 1980:91-92）。

* 本文是国家社科基金重大项目（21&ZD304）阶段性成果，国家社科基金重大项目（19ZDA300，12&ZD174）资助，国家自然科学基金“从世界语言透视东亚人群和语言的起源与演化”（31271337）项目资助。

布氏说：pin 跟 fin（鳍）、sin（罪恶）、tin（锡）以相同的声音收尾，可是开始的声音不同；布氏又说：pin 跟 pig（猪）、pill（药丸）、pit（坑）以相同的声音开始，不过收尾不同。布氏还说：pin 跟 pen（笔）、pan（盘子）、pun（双关语）的开头和收尾都相同，不过中间部分不相同。布氏最后说，pin 跟 dig（挖）、fish（鱼）、mill（磨坊）的开头和收尾都不同，不过中间部分是相同的。布氏总结道："一个词有三部分，我们就是这样改变其中任何部分，找到那些跟 pin 部分相同的形式"，"三者之中的每一个，都是区别性语音特征的最小单位，这就叫做一个音位"。美国结构主义的分析就是这样跳过了音节。他们不说 pin 跟 pig 是不同的（音节）单位，而是说它们是首音、音核或者尾音不同的音位。所以有评论说布氏从未定义音节也是不奇怪的（C.E. Cairns & E. Raimy 2011:12）。

实际上在 19 世纪末期以来有一些极富创新的研究从不同角度提出了音节单位问题。例如，Bell（1911）在《言语机器》论著中就关注音节现象，并推动音节在聋哑人教育中起了重要作用。还有令人诧异的是，Henry Sweet 编撰的《语音学手册》（1877）中，音节是语音学论述的主题，可是，Sweet 又跟 Paul Passy 一块研制出以元音和辅音音段为单元的国际音标。[①]这说明当时音节跟音段概念都是活跃在学术领域的重要概念。Stetson（1945）则转述 Sievers（1881）在《声音生理学基础》中的观点：他把辅音和元音分析为音节的功能，认为音节是语言的基本单位。这样的观点放在现代学术界差不多就是逆天了。

显然，到二十世纪初，音节已经迅速发展成为一个重要的描述性结构，产生了一个有趣且持久的分歧，Cairns & Raimy（2011:Pp8-9）回顾指出："我们是否应该把音段--或者，更好的是，把音段的序列看作是基本的，而把音节结构看作是次要的；或者音节是基本的，音段单独地根据其在音节内的功能来定义？"当时音节另一个主要观点是 Stetson（1945:50）提出的，即音节是以胸部收缩产生的声门下气压增加的脉冲，胸腔脉冲被认为是音节的语音对应物。在这样的认识基础上，Stetson（1945:11）指责 Bloch 和 Hockett 等美国结构主义者将语言学限制为一门分类科学。他说："言语声不是一系列分离的声音，就像相邻的串珠，一会儿这样产生，一会儿那样产生；它们是胸部脉搏、音节的相位，是基于呼吸装置的语音产生机制过程的一部分"。不过，60 年代以后，Ladefoged（1967）、Lieberman 等人（1967）报告说他们无法复制 Stetson 的发现。至于人们最重视的音节有内部结构观点是 Trubetzkoy（1939/1969）提出来的，后来 Pike 和 Pike（1947）丰富和发展了这套理论，附带提出了音节结构与重音和韵律的关系。以上描述包含了一个关键的模棱不清之处，那就是音节概念总是跟元辅音音段概念纠缠，（不知）哪个是必要概念，哪个是伴随概念。Cairns 和 Raimy（2011:13）评论说：与布拉格学派的信徒一样，结构主义者认为音节是强加在预先存在的音位串上的超音段结构。

生成语法问世之后，以 Chomsky 和 Halle（1968）的《英语的语音模式》为例可以清楚知道，生成语法从未认真考虑音节概念。现在回顾起来，这也算是从前人观点继承而来。换它的话来说，传统语言学或者结构语言学以来的体系都没把音节当回事，生成语法这儿也没它什么事。在音节问题上，生成语法作为新兴理论竟也没有比结构主义多迈出一步，它也不定义音节。直到 20 世纪 70 年代中叶以后，非线性音系学逐步兴起。该理论囊括了多种观念视角，形成自主音段音系学、CV 音系学、节律音系学等流派，但其中较为具有共性的一条是

[①] 音段概念一般即指音素，具体指元音(音段)或辅音（音段）。

重新认识音节，大幅提升音节地位，渲染音节的价值和作用。有了这样一股推力，音节真的成为一级实体单位了吗？其实是没有的。

吸引非线性音系学关注音节的主要因素是超音段特征附载于元音或者音节这个原因,以及由此引发的诸多难题。人们在分析中也透视出重音和声调等现象并非元音音段的固有特征，都强调超韵律的重音、声调、音长、音步和莫拉等特征可以通过音系分层分别研究。以 20 世纪 80 年代左右 McCarthy（1979）等人创建了 CV 音系学来看，该派强调音节在分析中的价值,能对音节内部结构和功能做出合适分析,使音系表述更具合理性。例如增加音节层：

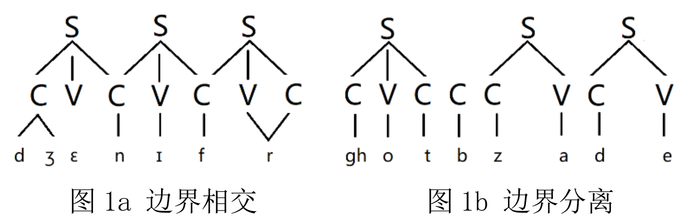


图 1a 边界相交　　　　图 1b 边界分离

但在本质上，这个所谓音节层只是 Kahn（1976）在音节层和音段层（元辅音）之间建立的一个称为 CV 的虚拟层（参看图 1）。其功能是为了更合理地解释音节上音系和音位对音节峰、音节边缘、音节轻重、元音长短、辅音丛和元音丛等概念的表述和技术处理。不过，该理论的主要建构者 Clements 和 Keyser（1983）提出一个非常高明核心音节的框架，呈现为音节结构四种类型：

A: CV　　　C: CVC

B: V　　　D: VC

这个结论是采用音系学规则推导出来，并通过世界各地语言的音系验证获得的。例如，如果一种语言具有 D 音节类型，它一定存在前三种类型。特别是 A 类型，是现存人类语言都有的类型，具有普遍共性，因此被称为核心音节结构。

再晚些时候优选论兴起，提出音节结构制约律。其中 Kager（1999:97）对世界语言音节的音节首音、音节尾音，即音节边界类型加以描述和讨论，其基本思路如表 1 所示：

表 1 音节边界复杂性的类型（黑体表示辅音丛边界）

	仅简单尾音	允许复杂尾音
仅简单首音	CV，CVC （日语，Yokuts 语）	CV，CVC，CV**CC** （芬兰语，Tunica 语）
允许复杂首音	CV，CVC，**CC**V，**CC**VC （西班牙语[①]，Sedang 语）	CV，CVC，CV**CC**，**CC**V，**CC**VC，**CC**V**CC** （英语，Totonac 语）

Kager（1999:93-98）根据音节结构实际案例提出四条制约规则，每两种音节之间的蕴含关系清晰。（此处依据李佳（2017:149）归纳。Kager 原文出处页码附后）：

表 2 Kager 音节结构制约律

	蕴含共性	制约律名	制约律形式表达	制约律内涵	
首音	无首音音节⊃有首音音节	ONSET	*[$_\sigma$ V	音节应有首音	93
尾音	闭音节⊃开音节	NO-CODA	*C]$_\sigma$	音节应无尾音	94

[①] 原注：西班牙语极少允许复杂尾音，例如 biceps(Harris 1983)

复杂首音	复杂首音⊃简单首音	*COMPLEX^{ONS}	*[$_{σ}$ CC	首音应简单	97
复杂尾音	复杂尾音⊃简单尾音	*COMPLEX^{COD}	*CC]$_{σ}$	尾音应简单	97

从传统语音学到布拉格学派，再到美国结构主义，音节仅仅作为一种描述术语填补人们对音系分析的叙述便利，就像花瓶一样，好看不中用。更普遍的情况是，人们往往是在词的语音形式下开展音位分析，分析对象直接就是构成词形的元音和辅音。即使是多音节词，人们也可以越过音节开展研究。这就是音节的命运，表面上它是介于音段层次和语素（或词）层次之间的一个层次，并且为音位变体提供了语素或者词的结构和边界语境，有利于分析。实质上，音节是可以略过的环节，没有任何机会充当语言学或语音学必有的一个或一级单位。

1.2 音节的价值

音节虽然长期处于语言分析的伴随或者陪衬位置，但是音节自身若隐若现的功能却一直促使人们深思它所蕴含的价值。

上文回顾了人们对音节的历史认知，不难发现非线性音系学的讨论全面拓宽了人们对音节的看法，很多以前模糊不清的观念逐步清晰起来。这种认识波及“系统”和“结构”两个层级。“系统”指的是就音节整体和音节与系统关系而言的，包括词（系统）；“结构”指的是音节内部分析而言的。

第一，语言学家普遍认为音节是音系单位之一，这意味着如果像美国结构主义那样完全略过音节直奔元辅音音段分析可能会错过某些重要语音属性。例如音系学发现重音和声调等韵律要素的分布都相当程度上涉及音节结构条件，这是因为音节是重音、声调等韵律现象的载体（王洪君 2008:94-97）。在这个意义上，音节的价值和作用不可能是可有可无的，甚至可以说，音节是各种韵律单位实现其功能的最基本领域。

第二，在有关音节的组织理论中，目前普遍接受的是 Jespersen 的响度说。因为响度说主要依据音节内部音段语音属性来组织，产生了响度顺序原则，即：塞音<擦音<鼻音<边近音<近音<元音。（王洪君 2008:99）尽管这个原则在世界语言中存在一定的反例，但总体上能反映出音节内部音段之间的音峰和边界序列关系，有利于人们对多音节词的音节划界操作。

第三，在音节结构分析上，人们提出按照内部音段性质切分和分层的方法，例如音节首、音节核、音节尾。进一步，借助世界语言的音节结构类型普遍性分析，人们发现一些重要的音节分布类型，例如南岛语系斐济语（Fijian）没有韵尾辅音，音节是典型的 CV、或(C)V 类型。（Dixon V1:278）这一现象与英语等欧洲语言绝然不同。那么，怎样的结构是更普遍的类型呢？可能受这类现象启迪，Kahn（1976）提出首音最大原则，即将词中辅音尽可能划归后一音节首，例如 VCV 应划分为 V.CV，而不是 VC.V。

第四，音节是词的外显形式，连接着载义问题。特别是多音节词的音节之间可能存在语音特征相关性，例如两个音节叠音或半叠音：叠音：tɕi^{24}tɕi^{213}“济济”，辅音半叠：pʰan^{35}sʰan^{55}“蹒跚”，元音半叠：kyagkyog 藏语“弯曲”。这类现象大多与词音节（参见第 5.3 节）及词义的产生和演变有关。（江荻 2013，2021，2022）因此，对音节的类型学认识很有可能反映出音节形式的历史渊源。音节与词义的极度关联问题有时也很凸显，这是所谓单音节性语言的

特征，例如汉语语音及相关分析无法跳出文字、音韵、训诂狭小的“声韵调”音节框架，这与音位理论“一味地把语音单位往小里切分，在理论上不具备完整性”算是两个极端。（陈保亚 2015:119）

但是，音节呈现出的这些价值讨论似乎主要围绕音节结构和功能而言，未阐释不同音节结构的性质差异，尚不能回答音节作为独立整体单位的价值，使得我们仍然无法定义音节[①]。

不定义音节能研究语音吗？从目前状况看，似乎学者们都持肯定的意见。这让我想起了端木三教授（Duanmu, 2009）提出的普遍性见解：

> “正如缺乏生命（或死亡）的定义并不妨碍生物学家研究生物，缺乏音节的定义也不应妨碍我们研究音节。我将论证，关于音节的许多问题都可以被解决，并得到合理的回答，比如最大的音节大小是什么，可能的音节首音是什么，以及如何确定音节的边界。这些结果构成了对音节理解的具体进展”。

2. 困惑与出路

2.1 皮特·拉迪福吉德

音节的本质和作用一直是语音学和音系学的痛点，不揭开这层迷雾，语言学的基石总是有点让人不放心，学术也难以进步。同时，即使随着分析技术的进步，有些研究能够更深入地逼近音节本质的研究，取得新的进展，却无法达到音节的真相。这样的状况对于那些理论悟者，究竟何为音节本质带来的困惑令他们无法解脱。

最典型的事件来自赫赫有名的语音学家皮特•拉迪福吉德教授（Peter Ladefoged）。他在《元音和辅音》（2012）这本著作的第16章(P187)指出：[②]

> 当我们说话的时候，我们真的是在把辅音和元音连接在一起吗？我认为答案是否定的。当我们说话的时候，我们并没有元音和辅音连接在一起，原因很简单，他们并没有分开存放。说话（行为）涉及从大脑某部分抽取词的存储形式，但是词不是以声音序列存储。他们是以整体形式存储的，或者至少以整体音节（形式）存储的，其中辅音和元音不是分离的单元。我们甚至应该考虑辅音和元音是否存在，书面上记录词的书写形式则可以另论。对于一个撰写《元音和辅音》著作的人来说，这个想法实在是一个很奇怪的思想。我会告诫你的是，很多研究言语的人是不会同意下面两节的观点的。但是我希望让你看到，辅音和元音很大程度上是我们良好科学想象力的虚构。

这段话简直就是警言，不可能不令人震撼。拉迪福吉德（1925-2006）何等专业人士？世界著名语音学家，曾任美国加利福尼亚大学洛杉矶分校语音实验室主任、国际语音学会会长、美国语言学会会长，被认为是20世纪贡献最显赫的语音学家之一。他在《元音和辅音》第16章（P199）结语中再次指出：

[①] 我们注意到中国学者呈现出理论完美主义倾向，有不少教材和论著曾尝试给音节下定义。例如罗常培、王均（1981）、叶蜚声、徐通锵（1981）、石锋、冉启斌（2019）等。

[②] 本书 2001 年初版，2012 年出版第 3 版，这部分内容从第 15 章改为第 16 章。由于新版增加 Disner 作为修订人，端木三教授审读本文初稿时建议采用第 1 版。鉴于第 3 版的中译本 2022 年底已经出版（赖福吉，本文译作拉迪福吉德，2022），为方便读者核查，本文仍依第 3 版引用。

我的观点（跟许多言语学家不一样）是，我们以更像音节的单位而不是单个的音素来组织我们的语言活动。如果不提音节，我们就无法解释为什么实际并没有cat被说成act这样的言语错误。字母表把音节看作由元音和辅音这样的分离片段组合而成，却只是历史上希腊人的一次发明。它是让人们能够把词记录下来的一个聪明发明，而不是发现词是由大大小小音段组成的。

这个认识太重要了，用“发明”（be invented）和“发现”（discovery）来揭示音节的本质属性，直指字母表的实用性质。而元音和辅音音段仅仅只是为了记录词和音节的文字发明，只是为了书写词和音节这种真实语言单位而发明的符号。

2.2 费尔迪南·德·索绪尔

对结构主义创始人索绪尔，关于音节的困惑又何尝不是如此。在批评（1）元音响度说（作为音节基础）和（2）音节重音学说时，索绪尔（2001:59）提出：

“我们的方法与上述（1）和（2）的区别是显而易见的：通过分析音节在音链中出现的情况，我们找到了不可还原的单元，即开音和闭音；然后通过组合这些单元，我们能够定义音节边界和元音峰值。现在我们知道了声学效应必须发生的生理条件。上面批评的理论遵循相反的路线：从孤立的语音种类中，理论的支持者假装推断出音节的边界和响音的位置。在给定的一系列音素中，一种发音可能比另一种发音更自然、更容易；但总的来说，在开口和闭合发音之间进行选择的可能性仍然存在，音节划分取决于选择而不是直接取决于语音种类”。

索绪尔的意思不难明白，一方面他指责响度等学说是从孤立的音段，例如元音，并结合它的出现环境来假设音节边界和音节核，另一方面他认为存在完全不同的音节，例如开音节和闭音节，不同音节自身已经划分音节的边界。这是完全不同的视角，前者是孤立的音段观念，任何变异动则归咎线性音段的影响；后者是整体观念，即要看音节之间造成内部结构差异的生理特征，也看到音节与音节之间呈现的边界声学特征。不过，索绪尔似乎并未就此展开进一步的研究，他已经看到音节呈现出元音或辅音的性质（能够定义音节边界和元音峰值），却并没有放弃最传统的观念：音节是由元音和辅音等构成的。

一个世纪之前德·索绪尔（1982:92）提出“关于音节区分理论的批判”，无论这些理论本身是否发展完善，目前学界是否普遍将音节理解为表现音段响度的不同元辅音序列，科学的音节界定还是令人迷茫不知。

对语音基本单位感到困惑的学者当然不止以上两位。Abercrombie（1967:38）曾就音素字母的来源表达过与拉迪福吉德（2012）相似的观点，他说：“基于音节的音段书写系统曾经发明过一次，但就是这么一次：希腊人的这个杰出发现给我们带来了不同于音节的字母书写系统。所有当代的字母书写系统（有很多）都来自于希腊的系统。而对于大多数语言来说，字母书写系统是可以设计出来的最经济、最实用的系统，但事实证明，它所基于的系统的分析对于大多数人来说太难独立完成”。读者应该注意到，Abercrombie认为字母书写系统（alphabetic system of writing）是设计出来的，非凡人所能。更多人则关注它用极少的符号构建无限词语的经济性和实用性创见，不再关心语言本身的真相。

朱晓农也描述过赵元任的困惑：“他对待音位的理论态度和实际做法相去甚远”，并指

出他那部备受西人推崇的《音位标音法的多能性》(1934/2002)与《中国话的文法》(1968/2002)趣旨完全相异,这部晚年的毕生名作用以描写汉语音系的基本单位竟不是音位概念而是声韵调及音节表。“为什么赵元任分析汉语音系时舍弃音位概念呢?为什么他从早期的音位理论奠基者、传播者,发展到晚年的实践中竟弃音位如敝屣?” 这是因为存在音节概念和声韵调单位。(2010:336)[①]

关于音节本质的认识,出路何在呢?我们认为应从相辅相成的正反两面来反思。正面说,音节是普通人能感知的语音单位,我们有任何理由不把它作为语音单位处理吗?回答应该是:没有!如此来说,音节不能作为一级真正的语音单位,或者人们不知道音节究竟是什么的原因不在音节自身,我们应把这个原因找出来。反面说,替代音节作为最小语音单位的音段,例如元音、辅音和其他喉音看上去也是人们可以感知的语音单元,例如英语词尾复数形式[z/s]等,藏语使动前缀[s-](grim 紧,sgrim 搓紧)。不过,迄今发现的纯辅音等特定音段被人们感知是有条件的,大多是在语法条件之下,是通过学习获知的(幼年习得时他们并不自知)。仅此也可能启发我们,带条件认识的事物属性有可能扭曲事物的本质。假如音节是最小不可再分的语音学或音系学单位,语言学和语音学将呈现怎样不同的格局?

3 字母发明陷阱

3.1 文字和语言之间

众所周知,文字编码方式和语言本身的结构之间存在着一种天然关系,即文字的表音性。可是,世界范围内,文字的发明不是一时一地之产物,尤其在传播应用中有些文字不知道经历了多少不同的人群和语言,这些人群对文字进行改造以适应自己的语言,以至于今天的人们对历史上的“表音性”已不知所表为何。据克里斯坦《文字的历史》(2019),文字的表音性主要指表音节和表音素,其文字分别称为音节文字和字母文字。[②]

什么文字表音节?什么文字表音素?换句话说,文字的表音形式是什么决定的?这是一个极具争议的问题,如果拿当今最知名的文字来对标和鉴别:汉字是典型(语素-)音节文字,英文是典型字母文字,它们在语音上的差别既有单音节性对多音节性、简单元辅音音段序列对复杂音段序列,还有大量形态音位变体差异和韵律差异。[③]那么,汉字与英文对比起来,表音性是不是可以直接归结为由具体语言的语音特征决定呢?

从全球文字发展历史来看,这似乎不是一个正确的提问方式。根据文字学家的观点,最古老的楔形文字始于公元前四千纪末的美索不达米亚平原,即“乌鲁克“(Uruk)象形文字

---

① 朱晓农教授极力提倡建立音节学分支。他认为“音节是一个语言的音系中最基本、最核心、最重要的结构单位,它由一个声母和一个韵母组成。语言中能够自由运用的最小的语音单位构成一个音节”(2010:322)。不过,他强调语音学和音系学的区分,想在音节和元辅音之间找平衡,甚至用上传统音韵学(汉语历史音系学)术语,最终目的是强调音节作为音系学单位的价值。

② 罗杰斯主编的《文字系统:语言学方法》提出还有一种短音节文字系统(moraic writing system),短音节是介于音素和音节之间的语音单位,一个音节包括音节头、音节核和音节尾,而一个短音节只包括音节头-音节核,或者只有音节尾。英语 cut 包含两个短音节:/kʌ/和/t/。中国传统韵书另有一种声韵表音系统。此处对这四种表音形式简单加以区别:音节代表 CV 或 CVC 整体,音素指 V 和/或者 C,短音节指 CV 或者带韵尾音节的 CV 和 C,声韵指 C 和 V,或者带韵尾音节的 C 和 VC。

③ 通常认为文字有四种类型,分别是:语素文字、语素-音节文字、音节文字和字母文字。裘锡圭(2013:18)认为,“语素-音节文字”这个名称对早期和晚期的汉字都适用。

泥板（贝瑟拉 2015），但真正成熟的楔形文字直到公元前2500-2000年才出现，主要记录苏美尔语。最初的苏美尔文楔形文字以表意字为基础，逐步发展为同时采用表意符号和音节符号来构造，其中有一个重要特征是苏美尔文几乎都是单音节字，这就意味着每个音节可能表达多个意思。例如：（克里斯坦 2019）

/dug/ 瓮，好，说（三项词义）

/lu/ 繁荣，人，麻烦（三项词义）

公元前2004年，苏美尔王朝灭亡，征服者阿卡德人建立的巴比伦帝国沿用了苏美尔楔形文字直到公元前1200年。阿卡德语所属闪米特语是多音节的，经过一些符号的调整改造，阿卡德语同样也能用楔形音节文字表达，同样能够使用表意符号和音节符号。

埃及圣书字跟楔形文字差不多是同时代发明的文字，大约是公元前3500左右。圣书字是具有很强图画性的但语音符号化的文字，其内部表现出辅音字母记音的特征，只是这种所谓的辅音字母实际上反映的却是音节性文字。在近东地区，音节文字传承了近3000年。在此过程中，随着不同族群的衰落和兴盛，还出现过其他多种音节文字系统，例如克里特象形文字、线形文字A、线形文字B、古塞浦路斯音节文字。此外，在相邻周边地区也出现一些音节文字，例如原始埃兰文（今伊朗西南部）、印度河谷文字（克里斯坦 2019）。

回到文字表音性问题上来，历史的案例值得重视：语言的单音节性和多音节性跟文字的表音没有必然联系，单音节的苏美尔语和多音节的阿卡德语都采用了音节文字表征。因此，拿汉字跟英文对着说证明不了二者表音性的差异来自语言声音结构的差异。迄今为止发现，世界所有自源文字都是音节文字，不论发明得早还是晚，除了上述文字，还有中国的彝文、美洲的玛雅文等，音节文字反映了人类语言的真实语音本质。另一方面，音节文字未必是惟一能代表语言的书写符号方案，不排除符号的实用性以其他方式曲折实现表声音的可能性。

字母文字是以音节文字为基础产生的。最早发现的字母文字文献出现在西奈半岛（公元前1500年），据克里斯坦，“原始西奈铭文的27个字母符号，其造型明显借用了象形文字。不过二者的音值不同，原始西奈字母的音值是通过截取该象形文字所表示的西闪米特语单词的第一个音值而形成。例如，表示‘牛头’（西闪米特语读为*ˊalepha*）的象形字在原始西奈字母中被用于转写辅音[ˊ]（声门塞音[ʔ]/起首元音），而表示“房子”（西闪米特语读为*beyt*）的象形字则被用于指代辅音B。同理，图案为波浪形指称“水”（西闪米特语读为*mayim*）的象形字则用来表示辅音M。”（克里斯坦 2019:288）公元前11世纪之后500余年中，地中海地区（古黎凡特）发展出多种字母文字，其中最著名的是腓尼基文。

字母文字通过截取音节首音的音值创造字母读音（现代称为音素），这种方法本身证明字母文字不能代表真实客观的语言发音，这是一种发明。特别是这种方法相应于语言普遍的/CV/语音结构一般总是截取音节首的辅音，很难直接截取单纯的元音。理论上说，字母文字所截取的辅音从来没有以纯粹的状态存在过，它本身也是抽象化的产物(克里斯坦 2019:43)。结果，大约到了公元前8世纪中叶，从腓尼基文改造而来的希腊文开始采用其他方法添加元音（音素）字母，例如改造腓尼基辅音字形转换为元音字母。至此，希腊人创造出相对完整的字母表系统，并逐渐成为后世西方所有字母文字的源头，这是以字母替代语言单位的逻辑陷阱。[①]

[①] 欧美部分西方学者（例如 R.Logan 2012）认为，字母文字开启了西方分析型科学范式，产生了抽象、

西方人是希腊字母以及后来的拉丁字母、斯拉夫等字母的发明者和使用者，很自然就占有了字母文字的解释权。他们用字母直接记录语言，不知不觉之中字母等于了语音（音素），字母替代了语音。换句话说，当字母成为独立书写符号，当字母成为编排字母表的单元，当字母成为记录语音的单位，字母的元音符号和辅音符号分类就变成语音的分类，元音和辅音就成为语音最基本单位，称为音素。一旦西方人把文字和语言混淆起来，把字母的应用视作对相应语言单位的使用，无形中就行使了他们的字母文字解释权和定义权，其他人是很难跟他们争辨（二者的差异）并说服之的。

现在问题清楚了，元音和辅音是来自文字系统的符号和概念，拉迪福吉德（2012）的疑惑得以解释：字母映射和对应到语言声音的的确确是我们的科学想象。在来源上，元音和辅音从来不是直接从语言提取或抽象出来的语音单位，而是依据文字应用和想象发明的符号，我们没有理由给二者画上等号。

字母表这个伟大的发明为西方2000多年的文明建功立业，不仅推动物质发明进步，而且缔造了无数精神成就。但是，文字毕竟是工具，是帮助人们表达和记录语言的工具，这是它产生和存在的目的，跟是否准确反映语言的语音性质和结构没有必然关系。

3.2 音节表和字母表

西方人借助字母创造了语言中对应文字的一种单位，即音素。为此，字母表成为音素文字系统的表征，也是世代族群学习母语文字的指南。迄今，世界所有音素文字都有自己的字母表，上百种文字的字母表形成了遍布世界的所谓字母表文化。从历史角度看，有人说“字母表是腓尼基人对人类做出的最重大贡献之一，是人类最伟大的发明。”（Hitti 1961:102）不过，这样的认知长远看究竟利于人类（知识体系）还是相反，目前人们似乎还无法回答。就本文论述来说，字母的发明有可能带来了认知的偏误，掩盖了语言的声音本质。

上文明确地指出一个事实，所有原生的文字都是音节文字。音节文字实际也有音节表，例如西方文字学巨擘I.J.Gelb在《文字研究》（1952）著作中列举了线性文字A（p93）、线形文字B（p94）部分音节表。图1是该著列出的古巴比伦辅音表单辅音符号（实际是音节表，Uniconsonantal signs of the old Babylonian syllabary,P109）。

表3 古巴比伦辅音音节表

a	i	e	ú				
ba	bi	be	bu	ab	ib	eb=ib	ub
da	di	de=di	du	ad	id	ed=id	ud
ga	gi	ge=gi?	gu	ag	ig	eg=ig	ug
ḫa	ḫi	ḫe=hi	ḫu	aḫ	iḫ=aḫ	eḫ=aḫ	úḫ
ia	ii=ia	ie=ia	iu=ia	—	—	—	—
ka	ki	ke=ki?	ku	ak=ag	ik=ig	ek=eg	uk=ug
la	li	le=li	lu	al	il	el	ul
ma	mi	me	mu	am	im	em=im	um
na	ni	ne	nu	an	in	en	un
pa	pí=bi	pé=bi	pu=bu	ap=ab	ip=ib	ep=eb	up=ub
qá=ga	qí=ki	qé=ke	qú=ku	aq=ag	iq=ig	eq=eg	uq=ug

分类和演绎逻辑。本文认为，在文字和语言关系本体认知上，字母并不能代表客观真实的语音单位，字母所代表的元辅音音段只是音节的特定呈现形式。

ra	ri	re=ri	ru	ar	ir	er=ir	ur
sa	si	se=?	su	ás =áš	is=iz	es=?	ús=uš
ša	ši	še	šu	aš	iš	eš	uš
ṣa=za	ṣi=zé	ṣe=zé	ṣú=zu	aṣ=az	iṣ=iz	eṣ=ez	uṣ=uz
ta	ti	te	tu	at=ad	it=id	et=ed	ut=ud
ṭa=da	ṭi=di	ṭe=te	ṭú= tu	aṭ=ad	iṭ=id	eṭ=ed	uṭ=ud
wa	wi=wa	we=wa	wu=wa	—	—	—	—
za	zi	zé	zu	az	iz	ez=iz	uz

早期文字大多包含各种功能符号。以埃及象形文字来说，有表意符号、表音符号、限定符号（表语义范畴或分类），因此，不同文字音节表的音节数量各语言差别很大。一般认为，如果一种文字的基本构成符号在30个左右，则属于字母文字；如果符号数量达到数十个以至上百个，则是音节文字；如果符号有数百个，就是语素-音节文字；再如果数量多达数千个，这种文字是语素文字。（克里斯坦 2019:237）古埃及文字符号数量是734个，苏美尔楔形文字（公元前2000年）598个，原始埃兰文约1500个；玛雅文约900-1200个；汉语普通话音节表大致为400个左右，汉语粤方言用汉字记下来有3067字，音节数则是667个（不计算声调）。（邓思颖 2015:58）

由此可见，把音节分析为元辅音音段是一种基于功利性和实用性目的的历史行为，这样的历史事件无所谓对与不对。但是，该事件的后果促动人们在这个基础上构建出不完全符合但相对逼近语言事实的体系，并且通过很多操作解读弥补该体系的理论欠缺（例如辅音文字一般能对应独立辅音音素），并最终产出一套看上去完备的语音学系统和音素文字字母体系。也许，当今的语音科学尚未发展到一定要突破这个体系的地步，且留待未来吧。

4 音节的本质和定义

4.1 发音产生什么样的单位?

人类发音器官直接产出的基本声音（单位）是什么？这是心理学研究的一种重要领域，称为言语产生。据心理学家研究，言语产生有两个阶段，词汇选择和词形编码（余林、舒华 2002，张清芳、杨玉芳 2005，张清芳 2008）。词汇选择指的是从心理词库获取相应的词及其形态规则，词形编码关注用怎样的声音单元把要表达的词转化为言语编码：音韵(编码)单元(phonological encoding)还是音素(编码)单元(phonetic encoding)。前者指的是音节，后者是音素或音位（音段）。

可惜的是，迄今为止，这个问题并没有获得明确的答案。让我们具体了解一下该领域的研究状况。上世纪80-90年代，心理学界关于言语产生主要形成两类理论体系。一是不连续两阶段模型(序列加工模型)，由Levelt et al(1991)、Levelt、Roelofs等人提出(1999,1997)。该模型认为言语产生的音韵编码单元是音素，音节是音素与词框架的联系环节，音素按序列从左至右插入音节框架。换句话说，音节不是心理词典的存储单元，音素序列的音节化是临场产生的。另一系统是Dell等人（1984，1986）提出的交互激活理论，该理论认为言语编码的单元包括音素、音节、音节的组成成分（辅音丛、元音丛等），并假设音节结构和框架存储于心理词典之中，言语编码时各种单元同时被激活。

更复杂而激烈的争论来自不同语言的实验结果，即不同语言发音结构特点是否对言语编码单元和编码方式产生重大影响？

表4　不同语言的编码单元实验对比

方法	语言	特征	编码单元	实验者
掩蔽启动范式	法语	音节界限清晰	音节	Ferrand 1996
掩蔽启动范式	英语	音节边界不清晰	音节/音段	Ferrand 1997
掩蔽启动范式	荷兰语	复杂条件	音段	Schiler 1998
掩蔽启动范式	英语	复杂条件	音段	Schiler 2000
内隐启动范式	英语，荷兰语	CVCV,CCV,……	音段	Levelt,Roelofs,Meyer 1999
图画-词汇干扰范式	汉语单音节	CV,CVC	音段无效用	张清芳 杨玉芳 2005
内隐启动范式	汉语双音节	CV-CV,CVC-CV	音节	张清芳 2008
图画-词汇干扰范式	粤语单音节	CV,CVC	音段	Wong & Chen 2009

以上研究常常呈现出混乱结果。例如Ferrand（1997）的研究涉及音节边界条件，如果一个词有清晰的初始边界，例如BALCONY，则产生明显的音节启动效应，如果音节边界模糊，例如BALANCE，则不产生音节启动效应。此外，实验还涉及词长效用、声调和重音的干扰等问题，还有实验方式的合适性选择。表1列出了其中主要方法：掩蔽启动范式(Masked priming paradigm)、重复启动范式（Repetition priming paradigm)、内隐启动范式（Implicit priming paradigm)、图画-词汇干扰范式（Picture word interference paradigm)。应该指出，实验者一般根据所用方式对实验结果进行推断，因此是受限的。Wong & Chen（2009）利用单音节汉字对应粤语不同音节结构就只能采用图画-词汇干扰范式，同时也把实验框在了音节内部音段上，在字等于单音节和字等于词困境下，结论就只能是字段编码。近期有研究对这样的设计是相当不认可的（岳源，张清芳 2015)。那么，是不是不同的语言特征一定要与不同的范式相结合才可能获得较好实验效果呢？大多数人对此持肯定的意见。

面对如此多的不同实验结果和争议，比如音节编码单元之于法语。有人提出法语作为罗曼斯语族语言，是音节定时语言，重音不敏感，处于词尾，词的边界清晰；而英语和荷兰语是日耳曼语族语言，是重音定时语言，对重音敏感，有大量边界模糊的音节（Levelt,al et. 1999)。实际上，重音或声调究竟是附着于音段的还是自主韵律要素，这本就有争议，那么它们是在声音加工的哪个阶段用顺序增添还是整体插入来编码呢（Shattuck-Hufnagel 1979)？当人们寻找各种造成实验结果相互矛盾或者没有定论时，他们是否考虑过实验的出发点本源就已经是错误的？还有人提出词的频率效用或音节的频率效用也可能对判断编码单元产生影响（周晓琳 庄捷，舒华 2001)。这些都需要更进一步的研究。

2011年出版的《音节研究指南》刊载了Cholin的一篇重要论文，“音节存在吗?——言语产生中音节单位提取的心理语言学证据”，该文一上来就在导言中列出言语产生研究和实验中的7大疑问：1.音节存在吗？2.音节是否在言语产生过程中代表功能上相关的单元？3.音节是否长时记忆中存储于词形的一部分？4.音节独立于词形存在吗？还是只在词形编码最终阶段进入的独立存储单元？5.音节在何种平面发挥作用？或者说音节作为音系学和/或语音学单位的心理语言学证据是什么？6.音节是否构成适用的发音单位？7.我们如何预设不

同编码层级上音节单位的相互作用和协调。

这些问题以音节产生和编码角度囊括了该领域几乎全部相关问题，文章通过回顾词产生理论，以及对比音节和音素编码假设的心理语言学证据，最后用不同语言中音节频率效应验证和讨论音节的功能。文章的依据是：因为只有存储单元才会出现频率效应，根据心理语言学的反应时间数据，就可论证音节作为言语产生中的功能相关单位的存在。文章实验结果强烈地指向这样一个结论：音节在词形编码的后期阶段发挥了作用，而且很可能是以单独储存的语音单位的形式。即使采用所谓边界不清晰的荷兰语和英语实验，音节也是单独储存的语音单位。所以音节频率效应为音节是(单独)存储单元的假设提供了强有力的证据。

如果说言语产生的编码单元问题跟具体语言特征相关，就涉及跨语言的现象。可是令人疑虑的是：从进化意义说，人的语音系统进化时间相当有限，保守地说，现代人的语言能力出现或许仅十数万年而已，各地形成语言词汇系统可能只有数万年。发音器官的神经-生理运动，例如口腔开合、舌体长短、舌头移动、喉部音姿、气流呼应，都需要生物级别的超长时间的进化。言语产生的发音器官编码单元属于人类语言能力，这种能力的形成是漫长进化过程形成的。而现代语言的语音结构差异只是百年或者千年分化造成，（理查德·利基 1995:97）怎么可能是由语言结构特征差异来决定言语产生的编码单元类型呢？

4.2 音姿与发音同步机制

Abercrombie（1967）是音节学说的重要创建者，这得益于他提出的用肺部气流机制来解释音节的理论。这一理论认为，当肺部气流机制发挥作用时，呼吸肌以大约每秒五次的速度交替收缩和放松，因此，空气以连续的小口气排出。每次收缩，连同由此产生的空气噗噗声，构成了一个音节的基础。因此，在持有这种理论的人看来，音节基本上是语言器官的运动，而不是言语声音的特征。他进一步指出：“对音节的分析主要是指对相关的发音、嗓声和鼻音运动的分析，这些运动将胸部脉冲‘塑造成’有声语言。这种复杂的运动被分割成连续的元素或片段。因此，每个部分本身都是复杂的，有各种言语器官对其产生的作用”。（Abercrombie 1967:34-41）

虽然Abercrombie的观点当时并未引起学界重视，但是，令人惊讶的是，20年后非线性音系学提出发音音系学（Articulatory phonology），其基本原理就是采用发音音姿（articulatory gesture）作为音系的基本单位。所谓音姿就是发音器官的运动，或者说是发音器官的收紧动作。据该理论创始人Browman和Goldstein（1986）的观点，收紧行为由6个发音器官承担：唇、舌尖、舌体、舌根、软腭和喉，收紧动作可以单个器官产生，也可以几个器官组合，产生的状态称为音姿，不同的音姿就是不同的语音。

更重要的是，音姿是从空间和时间来界定的单位。每个发音器官在确定位置上的收紧和收紧程度产生不同类型的音值，不同音姿之间存在动态协同关系，或者同步或者掩蔽，对音值的形成和变化均造成影响。Browman和Goldstein（1986，1989）还建立一个坐标轴来表现发音动作之间的动态时间关系，即所谓相位原则（phasing principles），这些原则规定了音姿的时空协调，其中Y轴对应发音动作，X轴对应发音时间。如果发音动作出现在相同相位，则同步发音，如果出现在反相位，则顺序发音。例如：

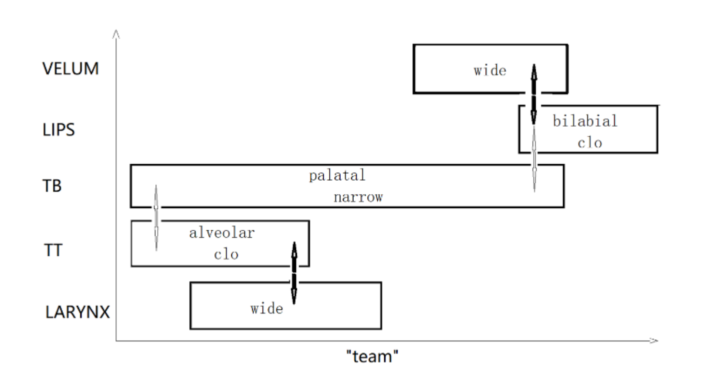


图2 音姿同步关系（注：VELUM软腭，LIPS唇，TB舌体，TT舌尖，LARYNX咽喉）

图2案例是英语词team（[tʰi:m]）。口腔器官运动过程由音姿描述，图2中的方框对应于单个音姿，发音启动时舌尖（起首齿龈辅音）部位收紧，咽腔张开，舌面（元音）收窄并与舌尖部位同步启动，舌尖持续后释放动作终止，舌面收窄释放的同时软腭下垂，随即双唇闭合收紧。这个过程完全对应[tʰi:m]发音过程。

发音音系学与形式音系学或者传统音系学差别较大，本文无意具体阐述。但有两点值得指出，一是该理论以音素为基础叙述音姿构成，二是音节剖析为与音姿对应的元辅音音段。由于发音音系学是基于生理运动视角，如果产生声学上的变化，该理论就无法加以解决。例如音段波形的瞬时性和叠置性，即使音姿不曾删减也可能会发生听觉上的掩蔽。再如，音姿重叠运动听觉上实际总是产生顺序声音，忽略了可能的同步声现象。发音音系学有关发音部位同步运动概念促使人们重新认识语音序列的时间结构关系（参看图3，Xu & Liu, 2006），并借助实验证明作为词的起始音节，辅音跟元音总是同时和同步的，即CV结构中C（initial C）与V同时发音，与发音音系学所说同相位同步发音一致。这项实验强力支持了语音单位的音节性质，也就是说，从发音角度看，不能把时间上同时发音并包含辅音和元音的音节割裂为辅音和元音两个单位。与发音音系学相应的另一种现象是：韵尾辅音结构(C)VC中的C（final C）处于反相位，只能接在前面CV之后顺序发音。[①]当然，这项研究还涉及更多复杂的现象，例如单纯元音起首的音节、辅音从起首的音节、多音节词内部音节之间的划界等等，该文及作者相关文章都有所论证（Xu, Y., F. Liu 2006; Xu, Y. 2020; Liu, Z., Y. Xu 2021; Liu, Z., Y. Xu and F. f. Hsieh 2022），有关细节本文不再赘述。

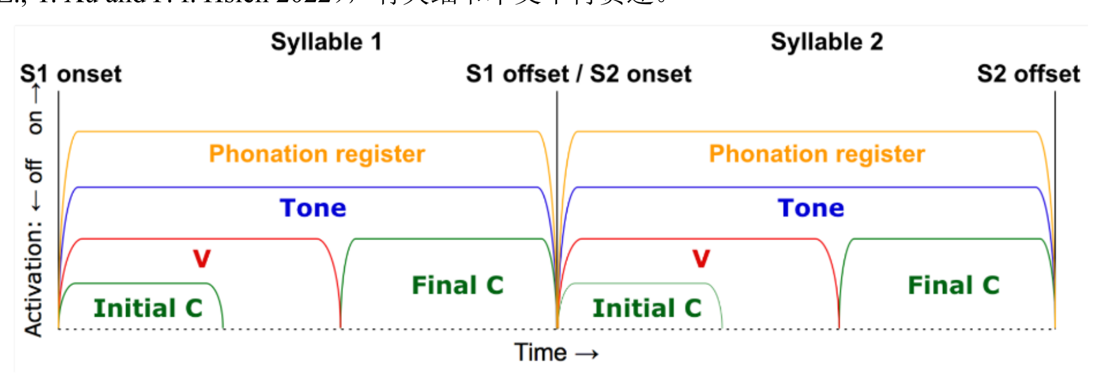


图 3 音节时间结构模型（Xu, Y., F. Liu 2013）

[①] 南开大学冉启斌教授审阅了本文初稿，对辅元音节同步发音完全赞同，且指出了 CV 为同一音节而 VC 感觉为两个音节。他的完整的音节感知观点发表在《音节的语音融合类型及其表现》（2012）等论文中，是对音节本质属性研究的重要贡献。后来他跟石锋教授合作发表的《音节的定义》（2019）融入了相关观点。

4.3 音节的定义

以上数节讨论环环相扣，逐步揭开了遮挡音节本质的纱幔，让人们有了再度审视音节定义的全新视角。在此基础上，我们尝试提出一个或许可以讨论的音节界定框架：

（1）人一次发音感知的声音单位，具有整体不可分割性

（2）人的语言最小可载义语音材料或组构载义用语音材料的单位

这个框架在内涵和外延上均需要做详尽解读，这是因为有不少原理性的内容隐含其间，例如同步发音原则、感知原则、音质内在性质、符号表征原则等。此外，由于很长时期以来，人们对音节的认知一直受到源自字母的元辅音音段的影响（详见5.1节），很多似是而非的概念杂糅其中，包括历史文字发明中的实用性或功利性认知，经历2000多年，其对字母文字使用者的深彻影响似乎已经难以扭转，并一定程度影响到非字母文字母语的研究者。

音节的定义首要是确定音节的性质，这是音节作为人类语言基本单位的核心问题。人们普遍知晓的音节概念源自听音层面上对语流的切分经验，特别是普通人的感知经验。可是，这种差不多很难出错的感知切分结果一直未能进入学术主流层面而概念化、术语化。因为，人们被另一种强大的认知吸引，即音节的多型形式：元音音节和辅音音节，[①] 可是，由于起始认知的偏误，这两种形式在诸多语用表征形式上竟又与文字体系中的元音字母和辅音字母相类。于是乎，这种来自字母表文字体系的单位携带着它构成音节的功能价值混入语音学系统，元音字母和辅音字母变身为语音学的元音音段和辅音音段，并自动获取了构成音节的功能，不经意间又将音节和元辅音音段分出等级。至今，语言实际使用者无法描绘出语言学意义上的元音是什么或者辅音是什么？而语言学家虽有困惑，也只能勉强把元辅音称作一种抽象的理论假设（拉迪福吉德 2012:187），还为此开始了漫长的元辅音音素的性质、功能、地位的论证过程，构建出所谓以元辅音为基本单位的语音学和音系学体系。[②]

回到音节的性质问题上来。音节是什么？音节是一段可载义或者组合起来载义的语音形式。这里的意思不难理解，共时上，音节是语言符号系统中词的外显形式编码，或者是词的直接物理语音形式，或者是构成词的物理语音形式音节段。来源上，音节专指人类进化过程中发展出一套以声音为载体的意义表征形式及其集合。如果撇开意义不论，一个音节就是说话人口腔发出来的一个声音，也是听音人感知到的一个物理声音（声波）。至于这个音节（声音）涉及到产生它的发音器官哪些复杂运动、包含了哪些复杂的物理波形，实际已在各种不同目的名义下开展了大量研究，只是未曾作为界定音节理论目标全面展开。

有了以上理解，我们继续讨论音节的定义。

1.按照初拟的框架，我们理解：音节是人类语言的声音现象，通过人体发音器官运动产生。这似乎排除了也曾经历漫长进化历程的其他灵长类生物或者其他哺乳动物，因为人类发音器官的进化达成直喉效应，使咽喉下沉拉长声道（孔江平 吴西榆 2021），又进化出合适

[①] “音节的多型形式”相当于说多样形式或变体形式，可比拟传统音位学说中音位的多个变体形式，各变体都是平等的音位。音节也有多个变体形式，此处为避免误解，采用多型形式指称多种不同类型的音节。

[②] 西方学者为什么未能厘清元辅音概念的来源，至少有两点可以推测，一是上文提及的被传承了近三千年强大的字母表文化对他们思想观念的挟持。二是很可能还有西方还原论哲学对文字符号分析的严重影响，就像语言学单位区分层级，词是由语素构成，音节也是由元辅音组成（音节也是文字单位）。实际上，还原方法对应着还原的意义和价值，词还原为语素不能脱离词义，音节还原为元辅音音段则很难表述其价值，特别是假定这个音节是从多音节词中取出的抽象音节。

的舌体长度便于在口腔形成各类可相互区别并分节的发音共鸣腔，并且产生气流控制和声门开闭控制机制。更重要的是，发音人和听音人通过语音范畴化系统获得神经生理和心理协调的进化，为产生音节奠定生理基础。

2.“一次发音感知的声音”表示同一发音个体有众多可对比的单次发出的声音，每次发出的音节不一定相同，或者不同发音个体之间有众多可对比的单次发出的声音，这说明每种语言系统都拥有相当数量的不同音节。历史上的一些古老书写系统音节表证明确实存在这样的语音单位。（I. J.Gelb 1952）

3. 所谓“（声音）整体不可分割”可以理解为从器官发音运动中切出一次发音行为，并体现为听觉上的单次声音（一次发音感知的声音），这也说明，音节作为人类器官发音和听音感知的单位具有一致性，这可能与习得相关。“整体不可分割性”还蕴含了音节的内在性质，无论发音器官是线性简单运动还是非线性复杂组合运动，无论物理声音是前后同质的还是变化的，例如同步从辅音性到元音性，还是同一音质贯彻始终，该声音都被感知为一次或一个整体声音，即具有同时性和同步性，并且是语感上能轻易感知到的声音单位。这又意味着音节是言语层面真实的、客观的声音，每个个体人都可发出和感知的音节声音。这是声音作为语言单位的逻辑基础。

4.“人的语言最小可载义语音材料”是一种科学描述，一方面叙述和建立这个概念，提出音节作为语言系统的基本单位，为之构建一个平台、一套术语和一个系统，另一方面挖掘它是一种形式载体，具有可载义的功能、应用范围，并能表征它的各种类型。

5.“组构载义用的语音材料的单位”反映了真实语言的面貌，指的是：音节是词的语音形式，词是不定长的。有些语言有更多单音节词，有些语言有更多多音节词，甚至有些语言完全没有单音节词。在这个意义上，音节相连会造成形式上的组合关系，或者相连音节怎样划界的操作，这种组合或分节关系在某些语言中可能带来一系列语用操作上的模糊认知或概念纠结，例如载义单音节词相连造成的复合词与组构载义用音节相连的多音节词有着本质的差别。

按照以上解读，音节可以按照以下表述明确定义，使其内容丰满、完善：

> 音节是人类个体发声器官同步同时运作发出的单次声音，这个声音是人类语言最小可载义或组合起来载义用的声音材料和语音单位或语音构造单位。

音节的定义不仅使语音单位的概念清晰起来，而且给语音学预设了一个新的研究路子，甚至是一个全新体系。

5 音节的性质和类型

5.1 元音是一种音节

理论上，音节作为人类语言最小声音单位，它自身是不能再度切分的。可是，由于字母文字发明的语用性优势导向（浅层理解为“先入为主”也未尝不可），西方语言学在一种模糊语音单位基础上业已建立宏大的语音学体系，该体系在语言和文字两套符号系统之中纠缠，虽然相当程度逼近语言系统，却永远不能达至理论完善，有时候甚至造成相互冲突或引向歧路。最关键的认知差异是元音与辅音，这是两个被认定为语音学最基本的音段单位，是

所谓构成音节的最小元素，也被认为是人类语言最小的基本符号，即元音音素和辅音音素。

元音和辅音真的是比音节更小的语音基本单位吗？音节是元音和辅音构成的吗？不回答这两个问题，就不可能真正理解音节，更不可能重新构建适合音节概念的多型元辅音音节的共存体系。事实上，第3节已经阐明元音音段和辅音音段出自文字上的字母概念，第4节阐明了音节的语音属性。可是，由于元音字母和辅音字母是对语音的映射，以及它们文字上构成音节的显见形式，使人们不知不觉一直混淆于两个相互关联而本质不同的音节概念之中，一个是文字拼写（处理）上的字母组合单位，一个是元辅音字母映射的真实语音单位。由此可见，音节术语原初的文字单位性质相当隐蔽，导致人们一直广泛将其作为语音本体单位对待。本文“音节”概念和术语是基于语音学意义的。

承接4.3节的“音节定义”，音节是发音器官产生的复杂语音单位。这个界定并不妨碍音节的元音呈现形式。换句话说，元音是发音器官能产生的音节形式之一。

实际上，如果人们希望从性质上界定元音，其定义大致是：元音是听感上的一种乐音。但是，这样的定义在认知上没有多大启迪价值，所以人们一般采用发音的和声学的方式来界定。例如拉迪福吉德（2012:26）说：元音是呼出气流不受任何阻碍产生的语音。更完整的说法是:元音是人类发音时不受齿、舌、唇阻碍并声带规律性振动而呼出气流产生的言语声。如果按照音节的定义方法，我们只能抽取元音定义的相关意思来表示，即，元音是发音器官单次产生的气流不受阻碍且声带振动的整体声音，（是人类语言最小可载义或组合起来载义用的声音材料和语音单位或语音构造单位）。把音节的定义这样套用于元音明显过于迂回，但我们的目的应该是达到了:元音也是一种音节，是本体性质上或类型上自成一类的音节，（物理声音）通常前后以自身起始或结束为边界，又分内部音质稳定或基本稳定和内部音质变动两类，前者称单元音，后者是复元音（假性的或真性的复合元音）。

5.2 辅音是不像音节的音节

再说辅音，辅音音段性质上一般被看做紊音，内部包含了湍流性质的物理声。此处，为简化讨论，我们有意略去辅音声学上那些复杂的描述，包括混波性、强频区、摩擦性、瞬流性、鼻音性、近音性等。辅音类别较多，如果阻碍或者摩擦仅仅来自口腔收窄处，可能是清辅音;如果阻碍来自口腔收窄处和声带，就形成浊辅音;鼻音略显特殊，一边放开鼻腔通道，一边造成口腔阻碍。总起来说，辅音是发音气流受到阻碍或摩擦产生的语音，有些可以延长，称为久音，有些不能延长，称为暂音或瞬音（王理嘉 1991:22）。

正是由于辅音的这种复杂现象，我们不可能简单套用音节的定义来界定辅音音节。而在真实语言中，普通人一般是不能切分出语流中的辅音音节的。如果有语言学家表示他们可以从语流中切分辅音音段，那也只能是为了某种功能目的编造的抽象概念，并不等于实际听到的语音，或者听到的只是某些特定的案例：所谓音节性的不完整辅音。例如粤语，韵尾[-m]是不除阻的鼻音，[-p˺]是所谓的无听感除阻塞音充当韵尾，实际不除阻。

元音可以定义为一种音节是因为元音音质决定的，称作元音音节（vocoid syllable），简称元音。[①]至于辅音是否具有音节性呢？完全可以通过一个程度阶梯量度逻辑判断，例如

① 当代语音学中作为音段的元音研究已具相当深度，甚至逼近其性质真值。可是对音段来源的合法性质疑不能不危及该学科的立论基础，当前可做的是，如果去除它的音段（音素）身份而赋予之音节身份，技术

擦音可以被认为是一种带无实体性元音的持续音节。一般认为塞音、塞擦音、擦音等不具有音节性，实际上放在全体音素队列中三者之间也存在音节性等级，以浊音为例，三者的音节性的顺序是：塞音 < 塞擦音 < 擦音，例如[d]<[dz]<[z]。阻塞音之外，其他辅音都一定程度上或一定条件下具有类似元音的音节性质，例如鼻音、颤音、边擦音等，近音最接近元音，通常具有音节性质。元音和辅音作为音段在已有的语音学体系中是一对平行概念，虽然辅音从性质上说不具有典型音节性，但在系统结构上或概念体系上却是具有元音音节同等地位的单位。这样的认知促使我们在构建新的音节概念中不能忽略它的价值，借助概念类推，我们尝试预设辅音也是发音器官产生的音节形式之一，即辅音音节。下文我们开展词音节性质和类型的讨论，并通过这个过程证实存在辅音音节（contoid syllable），简称辅音的理由。

5.3 音节的表示方法

上文已经指出，在传统语音学或现代音系学中，音节只是一个没有实际效用的花瓶，是描述元辅音音段与词（语素）之间关系的中介表述，或者更重要一点是描述声调-韵律作用时候的一个承载形式。一旦我们把音节作为实实在在的一个一级基本单位，我们至少需要一套描述音节的符号体系。这套体系能够反映音节的内部性质、音节的类型，以及音节的功能和作用。

继续讨论之前，本文要指出，不载义音节和音节类型实质上是抽象的和理论的表述，是从真实的词的语音形式提取出来的，尽管有些词本身就是单一音节。所以在指称词的语音或音节形式时候我们称之为词音节或词音节形式。此处我们要引入“音节结构偏好原则”概念（Lowenstamm J. 1996; Berces K. 2006; Levelt C. et al. 2000; Bonatti L L, et al. 2005; Capellan A R, et al. 2007; Nam H, Goldstein L, et al. 2009）。根据该原则，人类语音倾向音质上的辅元（辅+元，“+”号表示辅音和元音同步叠置，下同）组合音节，是一种复杂语音属性。结合上文4.2节发音同步机制和5.1元音音节和5.2辅音音节论述，加上这种内在性质从辅音性到元音性复合类型，至少我们已获得三种音节类型。[①]

江荻等（2014）提出音节（CV）或音节组合（CVC，VC，CV…CV）是人类语言基本形式，并转喻为人类语言“基因”，严格说则是“非生物性人类语言基因”。人类语言进化过程中，无论语音、词汇、形态、句法都可能发生演化，变得面目全非。但是有一种最基本的语音形式在人类语言中能够相当程度延续下来，即人类语言基因。我们曾用一种抽象化的公式描述这样的结构：{CV}p。CV表示辅音+元音类型音节，括号表示音节，括号外的{ }p表示不确定数量的音节（p=polysyllable）。例如{ }$_{p=2}$表示词形为双音节，{ }$_{p=3}$表示词形为三音节。{CV}p之所以被认为是非生物性人类语言基因，是因为它是惟人类拥有的符号系统建构形式。它是人类进化进程中借助人的生物性摄食器官衍生出的一套声音符号处理方法。人们需要做的就是观察、描写、记录这套方法。

从书写上说，最简易的音节单位表示方法是用S（syllable）表示音节，不同的音节可

处理上是否还存在不可逾越的系统障碍呢？这都是后续研究无法回避的问题。这些疑难对于辅音音段变身为辅音音节也同样存在。

① “音节结构偏好原则”概念是在音节由辅音和元音组合下提出的观点，现在我们不认为音节由辅音和元音组合，则不再提其内部存在“结构”，而是一种复杂属性。

以在引导符S之下添加数字方式来表现，例如S_1、S_2、……、S_x。但是，这个方法只能表示不同词的音节数量，不能揭示音节的性质和类型。

还有一种方法是在音节符号（引导符）之下列出音节内部可能的语音属性，例如：

符号Scv：表示音节内部具有不同性质的语音，即Ss。按照上文所述人类音节基本类型（Clements和Keyser 1983）或者“音节结构偏好原则”（Lowenstamm J. 1996），此处Ss的小写-s等于传统描述的CV，即Ss=Scv。为了叙述方便，下文以CV表征基本音节类型Ss，V表示又一种音节类型Sv,C表示第三种音节类型Sc。所以：

符号Scv ：表示音节内部是辅元复杂叠置发音性质

符号Sv ：表示音节内部是单纯元音性质

符号Sc ：表示音节内部是单纯辅音性质

不过，该方法的引导词S-实际是不必要的，就像非线性音系学中使用“σ”表征音节层级而无实质性作用。我们建议采用类似国际音标的表示方法，设计一个音节表达框架，本文采用双竖线“||”。例如，|S|可直接表示具有辅元同步发音性质的音节，即|CV|。

辅音符号C和元音符号V在音节属性中构成抽象的序列和平行关系，并且性质上相互对应。以逻辑关系论，如果我们抽取V作为一种音节|V|，那么剩余的C亦当作为一种音节，即|C|。|C|作为一种音节的假设有真实的经验基础，例如，car[kaːr]:|CV.C|。

|V|和|C|作为音节虽有一定客观真实支撑，但数量并不多，它们更可能是语言演化过程造成的音节变体，音节变体历时上包含形成的历史性。按照现代语音学的描述，还有一些常见的辅元音序列：CVC、CCV、VCC、CVV等，也应该是语言演变造成的音节变体。如何处理这类现象并纳入本文以音节为最小本体单位的语音新说，这是需要解决的问题。

综上，在体系的表达上，前述人类非生物性语言基因的表达{CV}p可以改为：|CV|p或|CV|p=*n*。对于音节|CV|*p*，我们可以将其中的C描述为音节的辅音属性或特征，V则描述为元音属性或特征，但不能将之割裂为|C|和|V|，音节|CV|是同时同步包含不同音质属性的单一复杂发音单位。音节|C|或|C·|和音节|V|或|·V|只能是历史演变造成的音节。[①]

5.4 划分音节的历史原则

当代语音学和音系学关于音节研究经常涉及的问题主要有音系结构及其相关的配列和层次分析，其次是音节划界及由此引起的音节边界和音节数量等诸多问题。关于音节结构，常见的分析是：音节首、音节核、音节尾，主要有以下三种分层和分段方法。

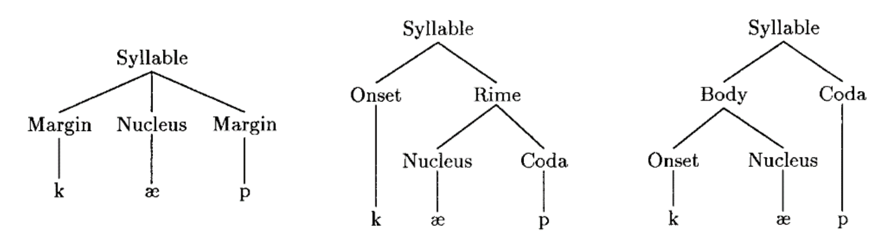


图4 当代音节结构分层分段分析案例（Kessler & Treiman 1997:297）。

音节划分更是理论峰出，提出各种方法和术语，诸如开音节、闭音节、音节峰、音节

[①] |C·|和|·V|表示基于历史演变产生的音节符号描述，而在共时状态下，可以直接表征为音节|C|和音节|V|。

边界和音节响度说，非线性音系学也提出音节首优先原则、核心音节模式、音节重量或作为韵律载体等分析概念。由于教学等语用目的，以上分析均带有很强的主观化方法，很可能给辅元音音段分析造成复杂性。本文拟从历史角度观察音节及其变体的来源，并以演化过程以及演化结果作为划分音节的依据。

上文曾提到"音节结构偏好原则"，实际这是从多学科和世界语言历史断面观察的结果，还包括了语音类型学推论。这个原则曾帮助我们建立人类语言初始状态的理想词形语音面貌假设，即|CV|p。在对$|CV|_{p=3}$的演化过程进行数学模拟计算时候，按照早前线性结构观点，我们也发现演化会导致大量突变形式，词音节数量逐步减少，原本规则的|CVCVCV|、|CVCV|、|CV|等词音节形式逐步复杂，产生了|CVC|、|VCC|、|CCV|、|CVV|、|VVC|、|CVCC|、|V|、|C|等多种变体词形。（江荻等 2014，Jiang D., *et al.* 2014）本案所谓突变专指元音或辅音属性的变异脱落，如果音节中的元音属性脱落，则称为负向偏移，系统中出现形式上的辅音连缀。如果音节中的辅音属性脱落，称为正向偏移，系统出现元音连缀形式。元音和辅音属性的连续脱落可能导致音节整体脱落和词音节长度的变化。偏移量和偏移角度是针对具体语言的词音节形式整体而言的。

此处不具体讨论计算过程。从演化的结果看，可用演化的参数，包括辅音和元音属性演变产生的偏移距离与偏移角来观察世界语言发生的变化。请观察表5。

表5. 179种欧洲、亚洲、大洋洲和非洲语言词音节演变抽样举例

序	语言名	区域	偏移量	偏移角	词长	序	语言名	区域	偏移量	偏移角	词长
1	English	Europe	1.2653	-0.2889	1.1304	9	Ge'ez	Africa	0.8256	-0.1747	2.0821
2	Dutch	Europe	1.5199	-0.2934	1.1942	10	Hadza	Africa	0.4521	-0.0153	2.2778
3	Norwegian	Europe	1.0118	-0.2167	1.4058	11	Lingala	Africa	0.4423	-0.0336	2.3744
4	Tamil	South Asia	0.8334	-0.072	2.1667	12	EasternFijian	Oceania	0.2116	0.0502	2.4497
5	Vietnamese	Southeast Asia	0.5076	-0.1265	1.058	13	Rapanui	Oceania	0.398	0.0995	2.7041
6	Thai	SoutheastAsia	0.686	-0.1935	1.1353	14	Malagasy	Oceania	0.3124	0.0415	2.8986
7	Beijing	East Asia	0.5577	-0.0507	1.2836	15	Maori	Australia	0.3333	0.0905	2.3137
8	Lhasa-Tibetan	East Asia	0.3508	-0.0668	1.6205	16	Japanese	Northeast Asia	0.4814	0.0857	2.4879

表5列出世界部分语言演变数据。欧洲、亚洲和非洲的语言偏移角为负，是元音属性脱落较多，形成辅音丛形式。大洋洲、澳洲以及靠近西太平洋的日语是辅音属性脱落多，产生元音丛。偏移量与词形长度有相关性，非洲和大洋洲语言辅元音属性脱落数量较为均衡，无论产生辅音丛还是元音丛都不算偏激，因此词长都保持在2音节之上。亚洲语言词长略高于一音节，偏移量不大，至少反映脱落结果辅元音属性保持均衡。泰米尔语属于南亚古老的达罗毗荼语系，词长接近太平洋语言。欧洲语言可能曾发生严重的音节内部演变，目前基本词汇词形较短。就总体样本而言，本案样本采用的是Swadesh-100词，核心词汇偏短，样本量也偏小，结论仅作为参考。

为直观的分析词形结构的偏移情况, 我们将单个词形结构的偏移数据映射到极坐标系当中的一个点上, 其极径与极角分别代表偏移距离与偏移角. 那么一个语言的偏移情况也就可以通过所有词形结构的偏移距离与偏移角来量度了（具体计算参看江荻等，2014。三角表东亚、东南亚，圆圈表欧洲、西亚、中亚、东北亚，星号表美洲、大洋洲、非洲）。

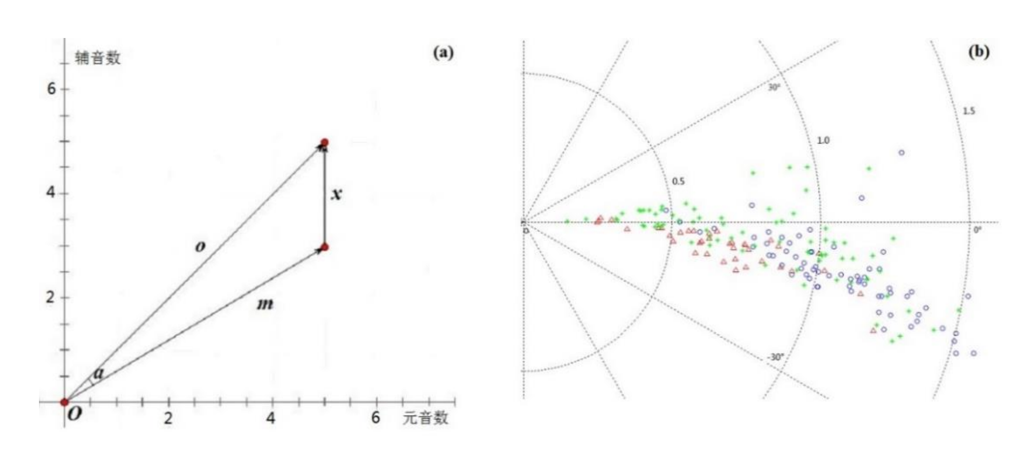


图4 世界179种语言词形演变造成的偏移状态

上举案例是语言词音节形式变化的结果，包括的词音节有各类复杂结构：|V|、|C|、|CV|、以及VC、CVC、CCV、CVV、CVCC等待分界形式。进一步，可以设立一定的原则将音节从词音节形式抽象出来：

1）基本音节形式|CV|不再切分，例如[pu]→|CV|，[tʰi]→|CV|，[kʷa]→|CV|。依附于辅音的送气和唇化等变体归于辅音，不单列。

2）首辅音优先，例如[tam]→|CV.C|，[ik]→|V.C|，ivy['aɪvi]→|V.CV|[①] 本类也可理解为排除辅音韵尾结构。

3）元音独立音节，例如:[ɛ]→|V|，[uə]→|V.V|。

4）复合元音不分割，例如[ai]→|V|,[əu]→|V|。如何区分复合元音与元音组合需建立规则。

5)辅音独立音节,例如[fri:]→|C.CV|,['rɪvə(r)]→|CV.CV.(C)|,[②]|ælps|→|V.C.C.C|。辅音作为独立音节除了5.2节给出的原因，还有一点也需指出，即|C|可以理解为带有无实体(substance-free)伴生元音(anaptytic vowel)的音节。

6）前附音节（一个半音节，sesquisyllabic）伴随元音的表示，例如达让语[pʰə³¹lɑŋ³⁵]→|CV.CV|“石头”，该词有时也记作首音节为次要音节的形式，其中起首辅音有点像复辅音，实际是带伴随元音的一个半音节：[pʰ·lɑŋ³⁵]→|L·.CV|。此处用字母L（light consonant）代表起首轻声或轻调音节的辅音（轻辅音），其中|L·|后随的伴生元音一般是偏央元音或与辅音同部位元音。这种现象在南亚语和部分藏缅语十分普遍，具有一定地域性。

7）附加特征和韵律成分暂不计入音节切分，包括送气、长音、鼻化、（硬/软）腭化、唇化、齿音化、重音、卷舌等。

当我们采用多型音节新观点来观察以上讨论，在三类音节中，|V|音节和|C|音节因其内在性质的基本一致性无需进一步解释，|CV|音节则需解读为是蕴含辅音和元音性质而同时同步产生的发音形式。所以，多型音节指的是：音节|S|呈现为|CV|、|V|和|C|三种音节类型，其中|S|也专指|CV|类型。考虑前附音节的多音节类型，则还应有|L·|和/或|LV|音节类型。

[①] 本文用英文句号“.”表示音节分界。例如|CV.C·|。多音节词不分界时候则连写，例如|CVC|。

[②] 在现有系统中，river 切分为 riv·er|CVC.V.C|，而不是*ri.ver|CV.CV.C|，有词义上的制约。本文基于历史演变原则分界，(The) Alps 切分是|V.C.C.C|，呈四个音节，与现有音节切分体系差别甚大。

6 结语

什么是音节？这是一个很长的故事，好几个世纪长了。本文综合前人观点并挖掘岐见所在来重新界定音节，以颠覆式方法提出元辅音性质新说和基于历史演化的音节划界，这些观点成与不成得看看以下问题怎么解决。音节上场作为语音学或音系学主角的话，诸如原本描述之元辅音的发音部位和发音方法何以待之，本文元辅音性质新说能否接受，还有与嗓音以及气流机制的关系，最重要的人们怎样重构以音节及其多型变体为基本单位的语音学整体版图，以及音节与词汇、形态、句法各系统之间的关系。总之，语音与字母分述（例如英语）、语音与汉字统论，都可能需要更多理论阐述和广泛的语用实践（参见裘锡圭 2013:9-20），譬如以元辅音字母为主角的代码、缩写、谜语、广告、游戏、排序、计算，这些语音之外的用途开辟了以语音为背景的字母应用新疆域，也值得拓展探索。

最后，作为结语和致敬，不能不提到令人疑惑和遗憾的一件事：拉迪福吉德既然提出了元辅音是人们科学想象的单位，他当年为什么不以实验来证之或证伪之呢？

致谢：本文初稿曾约请当代学术界多位语音相关著名学者审读，获益良多。在此特向以下学者致谢：中科院心理研究所杨玉芳教授、中国人民大学张清芳教授、密西根大学端木三教授、敦伦大学许毅教授、江苏师范大学朱晓农教授、中国社科院语言所覃远雄教授、南开大学石锋教授、李兵教授、冉启斌教授。杨玉芳教授等人依据汉语实验获得的音节作为言语产出单位结论就是不为西方旧学拘囿的创新；朱晓农教授先见之明地首创音节学概念并尝试理论上加以论证，这样的科学洞察力令人钦佩；石锋教授对音节本质的微观透视令人叹为观止，端木教授有关音节轻重价值观念富有启迪，催人思考。正是依托在这些贤俊们扎实的研究基础上，本文得以提出一些新的观点，并期待学界的批评指正。

主要参考文献

贝瑟拉，施曼特.丹尼斯 2015 《文字起源》，王乐洋（译），北京：商务印书馆。
布隆菲尔德 1980 《语言论》，袁家骅等（译），北京：商务印书馆。
陈保亚 2015 《20 世纪中国语言学方法论研究》，北京:商务印书馆。
邓思颖 2015 《粤语语法讲义》，香港：商务印书馆。
江 荻 2013 音节型语言演化的语音后果，《大江东去—王士元教授 80 岁贺寿文集》(Pp371-388)，香港:香港城市大学出版社。
江 荻 2022 藏语 ABB 式状貌词及其产生的理论根源，《语言科学》第 6 期第 634-647 页。
江 荻 2021 藏语 ABA'B 式状貌词与表现形态理论，《中国语文》第 6 期第 713-723 页。
江 荻，康才畯，燕海雄 2014 词形结构进化与世界语言的多样性，《科学通报》第 59 卷(21)：2084-2090。
克里斯坦，安娜·玛丽（主编） 2019 《文字的历史》，王东亮等（译），北京：商务印书馆。
孔江平，吴西榆 2021 人类语音起源和演化研究综述，《中国语音学报》第 1 期。
赖福吉 2022 《元音与辅音》，衣莉，兰婧晰（译），朱晓农（审定），北京:商务印书馆。
利 基 1995 《人类的起源》，吴汝康，吴新智、林龙圣（译），上海:上海科学技术出版社。
李 佳 2017 《音节结构与方言分组鄂东南方言的个案研究》，北京:中华书局。
罗常培，王均 1981 《普通语音学纲要》，北京:商务印书馆。

罗杰斯,亨利（主编） 2016 《文字系统：语言学的方法》，孙亚楠（译），北京:商务印书馆。
洛 根 2012 《字母表效应：拼音文字与西方文明》，何道宽（译），上海:复旦大学出版社。
裘锡圭 2013 《文字学概要》第二版，北京:商务印书馆。
冉启斌 2012 音节的语音融合类型及其表现，《汉语语音新探》，中国社会科学出版社，第 131-142 页。
冉启斌 2012 音素结合的融合程度与汉语语音的若干重要表现，《汉语语音新探》，中国社会科学出版社，第 143-169 页。
石 锋，冉启斌 2019 音节的定义--基于语言学的思考，《南开语言学刊》第 2 期 1-9 页。
索绪尔，费尔迪南·德 1980 《普通语言学教程》，高名凯（译），北京：商务印书馆。
王洪君 2008 《汉语非线性音系学》，北京：北京大学出版社。
王理嘉 1991 《音系学基础》，北京：语文出版社。
叶蜚声，徐通锵 1981 《语言学纲要》，北京：北京大学出版社。
余 林，舒华 2002 西方语言产生研究中的几个主要问题，《心理科学进展》第 3 期:248—255 页.
岳 源，张清芳 2015 汉语口语产生中音节和音段的促进和抑制效应,《心理学报》第 47 卷第 3 期:319-328 页。
张清芳 2008 汉语单音节和双音节词汇产生中的音韵编码过程:内隐启动范式研究，《心理学报》第 *40 卷第* 3 期:253-262 页。
张清芳，杨玉芳 2005 汉语单音节词汇产生中音韵编码的单元，《心理科学》第 28 卷第 2 期:374-378 页。
赵元任 1934/2002 《音位标音法的多能性》（叶蜚声译），《赵元任语言学论文选》750-795 页，北京:商务印书馆。
赵元任 1968/2002 《中国话的文法》（增订版），丁邦新译，香港:香港中文大学出版社。
周晓林，庄捷，舒华 2001 言语产生研究的理论框架，《心理科学》第 24 卷:262－265 页。
朱晓农 2010 《语音学》，北京:商务印书馆。
Abercrombie, D. 1967. *Elements Of General Phonetics*. Edinburgh University Press.
Bell, A. G. 1911. *The Mechanics of Speech*. New York: Funk and Wagnalls.
Berces, K. 2006. What's wrong with vowel-initial syllables? *SOAS working papers in Linguistics*, 14: 15-21.
Bonatti, L. L., Penña M, Nespor M, et al. 2005. Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, 16: 451–459.
Browman, C. P., L. Goldstein 1989. Articulatory Gestures as Phonological Units. *Phonology*, Vol. 6, No. 2, pp. 201-251.
Cairns, C. E., E. Raimy. 2011. *Handbook of the Syllable*. Brill.
Capellan, A. R., Schwartz J. L., Laboissière R, et al. 2007. Two CV syllables for one pointing gesture as an optimal ratio for jaw-arm coordination in a deictic task: A preliminary study. *Proceedings of the European Cognitive Science Conference*. 608-613. Delphes, Grèce. EuroCogSci.
Cholin, J. 2011. Do Syllables Exist? Psycholinguistic Evidence for the Retrieval of Syllabic Units in Speech Production. In: Cairns, Charles E., and Eric Raimy. *Handbook of the Syllable*. 225-254.Brill.
Chomsky, N. & Morris Halle 1991. *The Sound Pattern of English*. The MIT Press.
Clements, G. N., Samuel J. K. 1983. *CV phonology: A generative theory of the syllable*. Cambridge,

MA: MIT Press.
Dell, G S. 1986. A spreading activation theory of retrieval in sentence production. *Psychological Review*. 93(3)： 283-321.
Dell, G. S. 1984. Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(2), 222–233.
Dixon, R.M.W. 2010. *Basic Linguistic Theory*. Vol.1: Methodology. Oxford University Press.
Duanmu, S. 2009. *Syllable structure: The limits of variation.* Oxford: Oxford University Press.
Ferrand, L., Segui J., Garinger J. 1996. Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language*. 35:708-723.
Ferrand, L., Segui J., Humphreys G.W. 1997. The syllable's role in word naming. *Memory and Cognition*. 25:458-470.
Gelb, I.J. 1963. *A Study of Writing*. Chicago, IL: University of Chicago Press.
Goldstein, L., Browman, C.P. 1986. Representation of voicing contrasts using articulatory gestures. *Journal of Phonetics, 14,*339-342.
Hitti, P. 1961. *The Near East in History*. Princeton, NJ: Princeton University Press.
Jiang, D., Kang, C.J., Yan, H.X. 2014. Evolution of word-syllable structures and the diversity of world languages. *Chinese Science Bulletin*. Vol. 59, Issue 26, Pp3362-3368.
Jespersen, O. 1904. *Lehrbuch der phonetik*. (Translated by Hermann Davidsen.) Leipzig and Berlin: B. G. Teubner.
Jespersen, O. 1909. *A Modern English Grammar on Historical Principles*. Carl Winter, Heidelberg.
Jespersen, O. 1922. *A Modern English grammar on historical principles. Part 1: Sounds and spelling*. 3rd edn. Heidelberg: Carl Winter's Universitätsbuchhandlung.
Kager, R. 1999. *Optimality Theory*, Cambridge University Press.
Kahn, D. 1976. *Syllable-based generalizations in English phonology*. PhD diss., MIT.
Kessler B., Treiman R. 1997. Syllable Structure and the Distribution of Phonemes in English Syllables. *Journal of Memory and Language*. Vol.37:295-311.
Ladefoged, P. 1967. Three areas of experimental phonetics. Oxford: Oxford University Press.
Ladefoged, P. 2012. *Vowels and Consonants*(Third Edition, revised by S.F.Disner). Wiley-Blackwell.
Levelt C., Schiller N., Levelt W. 2000. The acquisition of syllable types. *Language Acquisition*, 8: 237-264.
Levelt, W., Roelofs, A., Meyer, A. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences.* 22:1-75.
Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., Havinga, J. 1991. The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122-142.
Lieberman, P., J.D. Griffiths, J.M., and R. Knudson. 1967. Absence of syllabic "chest pulses". *Journal of the Acoustical Society of America,* 41:1614.
Liu, Z., Xu, Y. 2021. Segmental Alignment of English Syllables with Singleton and Cluster Onsets. *INTERSPEECH.* Brno, Czechia.
Liu, Z., Y. Xu, F.f. Hsieh 2022. Coarticulation as synchronised CV co-onset: Parallel evidence from articulation and acoustics. *Journal of Phonetics*. 90:101116.
Lowenstamm, J. 1996. CV as the only syllable type. In: Durand J, Laks B (ed) *Current Trends*

in Phonology: Models and Methods: 419–442. Salford(UK): European Studies Research Institute, University of Salford.
McCarthy, J. 1979. On stress and syllabification. *Linguistic Inquiry* 10: 443–465.
Nam, H., Goldstein L., Saltzman E. 2009. Self-organization of syllable structure: a coupled oscillator model. In: Pellegrino F, Marsico E, Chitoran I, et al (ed) *Approaches to Phonological Complexity*: 299-328. Berlin & New York: Mouton de Gruyter.
Pike, K.L., E. Pike. 1947. Immediate constituents of Mazateco syllables. *International Journal of American Linguistics* 13: 78–91.
Roelofs, A. 1997. The WEAVER model of word form encoding in speech production. *Cognition.* 65:249-284.
Saussure, F.de 1959. *Course in general Linguistics*. Translated by Wade Baskin. New York: Philosophical Library.
Schiller, N.O. 1998. The effect of visually masked syllable primes on the naming latencies of words and pictures. *Journal of Memory and Language*, 39:84-507.
Schiller, N.O. 2000. Single word production in English: The role of subsyllabic units during phonological encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 26(2):512-528.
Selkirk, E. 1984. On the Major class features and syllable theory. In M. Aronoff & R. Oehrle, eds., *Language sound structure*. Cambridge,Ma: The MIT Press. Pp107-36.
Shattuck-Hufnagel, Stefanie. 1979. Speech errors as evidence for a serial ordering mechanism in sentence production. In *Sentence Processing: Psycholingistic studies presented to Merrill Garrett*, edited by William E. Cooper and Edward C. T. Walker. Hillsdale, 295–342. N.J.: Lawrence Erlbaum.
Sievers, E. 1881. Grundzüge der Phonetik: Zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen. Bibliothek indogermanischer Grammatiken. 2nd ed. Vol. 1. Leipzig: Breitkopf und Hartel.（Quote from Stetson 1945）
Stetson, R.H. 1945. *Bases of Phonology*. Oberlin, Ohio: Oberlin College.
Sweet, H. 1877. *A handbook of phonetics*. Oxford: Clarendon Press.
Trubetzkoy, N.S. 1969 *Principles of Phonology*. Foreword by Christiane A. M. Baltaxe, trans. Berkeley & London: University of California Press. [Original title: *Grundziige der Phonologie*, 1958. Third ed., 1962.1.]
Wong, A.W., Chen, H.C. 2009. What are effective phonological units in Cantonese spoken word planning? *Psychonomic Bulletin & Review, 16*(5), 888–892.
Xu, Y. 2020. Syllable is a synchronization mechanism that makes human speech possible. PsyArXiv. pp. 1–44.
Xu, Y., Liu, F. 2006. Tonal alignment, syllable structure and coarticulation: Toward an integrated model. *Italian Journal of Linguistics* 18: 125-159.
Xu, Y., Liu, F. 2013. Advances in prosody research and how they are related to other areas. In J. Wang and D. Chen (eds.), *Linguistics*, Beijing: Chinese People's University Press.

The Essence of syllables and A new explanation of the nature of vowels and consonants as syllables

Di Jiang
School of Language Sciences and Arts, Jiangsu Normal University, China
Institute of Ethnology & Anthropology, Chinese Academy of Social Sciences

Abstract: The syllable has long been regarded as a one-order phonetic unit without being known as an illusion; vowels and consonants have been the most solid units of phonetics without being known as impostors of letters; and the historical misconception that vowels and consonants make up syllables has gone unrecognized. The article analyzes the reasons why syllables have no linguistic status, describes the confusion over the unknown origin of vowels and consonants, analyzes the relationship between syllables and letters from a chronological perspective, and experimentally explores the unitary form of speech production, and reveals that the temporal structure of articulation prescribes the nature of syllables from a co-temporal perspective. Based on this, the paper redefines syllables in a subversive way, and reshapes the true status and value of vowels and consonants. The article concludes by suggesting that syllables evolved in a long history and eventually produced complex word syllable structures and limited pure syllable forms through shedding, whereby a set of minimal syllable analysis schemes and syncopation principles can be proposed.

Key words: syllable definition, vowel syllable, consonant syllable, alphabet letter, Syllabic segmentation